



1600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/581,976B

DATE: 11/26/2002

TIME: 09:26:10

Input Set : A:\seqlist.txt  
 Output Set: N:\CRF4\11262002\I581976B.raw

4 <110> APPLICANT: Dalemans, Wilfried L.J.  
 5 Gerard, Catherine Marie Ghislaine  
 7 <120> TITLE OF INVENTION: Vaccine  
 10 <130> FILE REFERENCE: B45124  
 12 <140> CURRENT APPLICATION NUMBER: 09/581,976B  
 13 <141> CURRENT FILING DATE: 2000-06-20  
 15 <150> PRIOR APPLICATION NUMBER: PCT/EP98/08563  
 16 <151> PRIOR FILING DATE: 1998-12-18  
 18 <150> PRIOR APPLICATION NUMBER: GB 9727262.9  
 19 <151> PRIOR FILING DATE: 1997-12-24  
 21 <160> NUMBER OF SEQ ID NOS: 28  
 23 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 220  
 27 <212> TYPE: PRT  
 28 <213> ORGANISM: Artificial Sequence  
 30 <220> FEATURE:  
 31 <223> OTHER INFORMATION: Chimaeric protein (protein D from Haemophilus  
 32 influenza B and E7 from Human papilloma virus type  
 33 16)  
 35 <400> SEQUENCE: 1  
 36 Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
 37 1 5 10 15  
 38 Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
 39 20 25 30  
 40 Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
 41 35 40 45  
 42 Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
 43 50 55 60  
 44 Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
 45 65 70 75 80  
 46 Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
 47 85 90 95  
 48 Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
 49 100 105 110  
 50 Ala Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu  
 51 115 120 125  
 52 Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser  
 53 130 135 140  
 54 Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro  
 55 145 150 155 160  
 56 Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser  
 57 165 170 175

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/581,976B

DATE: 11/26/2002  
TIME: 09:26:10

Input Set : A:\seqlist.txt  
Output Set: N:\CRF4\11262002\I581976B.raw

58 Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu  
59 180 185 190  
60 Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser  
61 195 200 205  
62 Gln Lys Pro Thr Ser Gly His His His His His His  
63 210 215 220  
65 <210> SEQ ID NO: 2  
66 <211> LENGTH: 663  
67 <212> TYPE: DNA  
68 <213> ORGANISM: Artificial Sequence  
70 <220> FEATURE:  
71 <223> OTHER INFORMATION: Chimaeric protein (protein D from Haemophilus  
72 influenza B and E7 from Human papilloma virus type  
73 16)  
75 <400> SEQUENCE: 2  
76 atggatccaa gcagccattc atcaaataatg gcgaatacc 60  
77 attattgctc accgtgggtgc tagcggttat ttaccagagc atacgttaga 120  
78 cttgcgtttg cacaacaggc tgattattta gagcaagatt tagcaatgac 180  
79 cgtttagtgg ttattcacga tcactttta gatggcttga ctgatgtgc 240  
80 ccacatcgtc atcgtaaaga tggccgttac tatgtcatcg actttaccc 300  
81 caaagttag aaatgacaga aaactttgaa accatggca tgcattggaga 360  
82 ttgcattaaat atatgttaga ttgcacca gagacaactg atctctactg 420  
83 ttaaatgaca gctcagagga ggaggatgaa atagatggc cagctggaca 480  
84 gacagagccc attacaatat tgtaacctt tggcaagt gtgactctac 540  
85 tgcgtacaaa gcacacacgt agacattcgt actttggaa acctgttaat 600  
86 ggaattgtgt gccccatctg ttctcagaaaa ccaacttagt gccaccatca 660  
87 taa 663  
89 <210> SEQ ID NO: 3  
90 <211> LENGTH: 822  
91 <212> TYPE: DNA  
92 <213> ORGANISM: Artificial Sequence  
94 <220> FEATURE:  
95 <223> OTHER INFORMATION: Chimaeric protein (protein D from Haemophilus  
96 influenza B and E6 from Human papilloma virus type  
97 16)  
99 <400> SEQUENCE: 3  
100 atggatccaa gcagccattc atcaaataatg gcgaatacc 60  
101 attattgctc accgtgggtgc tagcggttat ttaccagagc atacgttaga 120  
102 cttgcgtttg cacaacaggc tgattattta gagcaagatt tagcaatgac 180  
103 cgtttagtgg ttattcacga tcactttta gatggcttga ctgatgtgc 240  
104 ccacatcgtc atcgtaaaga tggccgttac tatgtcatcg actttaccc 300  
105 caaagttag aaatgacaga aaactttgaa accatggca tggccgg 360  
106 cgaccccgaa agttaccaca gttatgcaca gagctgcaaa caactataca 420  
107 tttagaatgtg tggactgcaa gcaacagttt ctgcgacgtg aggttatatga 480  
108 cgggattttt gcatagtata tagagatgg aatccatgt ctgtatgtg 540  
109 aagttttattt ctaaaattttt tgatgtataga cattattttt atagttgtt 600  
110 tttagaacaggc aataacaacaa accgttggtt gattttttt ttaggttta 660  
111 aagccactgt gtcctgaaga aaagcaaaga catctggaca aaaagcaaag 720  
112 ataagggggtc ggtggaccgg tcgatgtatg tcttggca gatcatcaag 780

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/581,976B

DATE: 11/26/2002

TIME: 09:26:10

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I581976B.raw

113 gaaaccacg c t g a c t a g t g g c c a c c a t c a c c a t t a a 822  
116 <210> SEQ ID NO: 4  
117 <211> LENGTH: 273  
118 <212> TYPE: PRT  
119 <213> ORGANISM: Artificial Sequence  
121 <220> FEATURE:  
122 <223> OTHER INFORMATION: Chimaeric protein (protein D from Haemophilus  
123 influenza B and E6 from Human papilloma virus type  
124 16)  
126 <400> SEQUENCE: 4  
127 Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
128 1 5 10 15  
129 Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
130 20 25 30  
131 Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
132 35 40 45  
133 Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
134 50 55 60  
135 Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
136 65 70 75 80  
137 Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
138 85 90 95  
139 Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
140 100 105 110  
141 Ala Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu  
142 115 120 125  
143 Cys Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val  
144 130 135 140  
145 Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe  
146 145 150 155 160  
147 Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Cys  
148 165 170 175  
149 Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr  
150 180 185 190  
151 Cys Tyr Ser Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro  
152 195 200 205  
153 Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys Pro Leu Cys  
154 210 215 220  
155 Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn  
156 225 230 235 240  
157 Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser  
158 245 250 255  
159 Arg Thr Arg Arg Glu Thr Gln Leu Thr Ser Gly His His His His His  
160 260 265 270  
161 His  
164 <210> SEQ ID NO: 5  
165 <211> LENGTH: 1116  
166 <212> TYPE: DNA  
167 <213> ORGANISM: Artificial Sequence

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/581,976B

DATE: 11/26/2002  
TIME: 09:26:10

Input Set : A:\seqlist.txt  
Output Set: N:\CRF4\11262002\I581976B.raw

169 <220> FEATURE:  
170 <223> OTHER INFORMATION: Chimaeric protein (protein D from Haemophilus  
171 influenza B and E6E7 fusion from Human papilloma  
172 virus type 16)  
174 <400> SEQUENCE: 5  
175 atggatccaa gcagccattc atcaaataatg gcgaatacc accatggaaatc agacaaaatc 60  
176 attattgctc accgtgggtgc tagcggttat ttaccagagc atacgttaga atctaaagca 120  
177 cttgcgttgc cacaacaggc tgattattta gagcaagatt tagcaatgac taaggatgg 180  
178 cgttttagtgg ttattcacga tcactttta gatggcttgc ctgatgttc gaaaaaaattc 240  
179 ccacatcgtc atcgtaaaga tggccgttac tatgtcatcg actttacctt aaaagaaatt 300  
180 caaagtttag aaatgacaga aaactttgaa accatggcca tggttcagga cccacaggag 360  
181 cgaccaggaa agttaccaca gttatgcaca gagctgcaaa caactataca tgatataata 420  
182 ttagaatgtg tgtactgca gcaacagttt ctgcgacgtg aggtatataa ctttgcttt 480  
183 cgggatttat gcatagtata tagagatggg aatccatataa ctgtatgtga taaatgttta 540  
184 aagttttattt ctaaaaatttt tgagttataga cattattgtt atagttgtt tgaacaaca 600  
185 ttagaacacgc aataacaacaa accgttgcgtt gattgtttaa ttaggtgtat taactgtcaa 660  
186 aagccactgt gtcctgaaga aaagcaaaaga catctggaca aaaagcaaaat attccataat 720  
187 ataagggttc ggtggaccgg tcgatgtatg tcttgcgtca gatcatcaag aacacgtaga 780  
188 gaaacccaggc tgatgcattt agatacacctt acattgcattt aatataatgtt agatttgc 840  
189 ccagagacaa ctgatctcta ctgttatgaa caattaaatg acagctcaga ggaggaggat 900  
190 gaaatagatg gtccagctgg acaaggcaga ccggacacagag cccattacaa tattgttaacc 960  
191 ttttgcgtca agtgcgtactc tacgcttcgg ttgtgcgtac aaagcacaca cgtagacatt 1020  
192 cgtactttgg aagacactgtt aatgggcaca ctaggaattt tttttttttt ctgttctcag 1080  
193 aaaccaacta gtggccacca tcaccatcac cattaa 1116  
195 <210> SEQ ID NO: 6  
196 <211> LENGTH: 371  
197 <212> TYPE: PRT  
198 <213> ORGANISM: Artificial Sequence  
200 <220> FEATURE:  
201 <223> OTHER INFORMATION: Chimaeric protein (protein D from Haemophilus  
202 influenza B and E6E7 fusion from Human papilloma  
203 virus type 16)  
205 <400> SEQUENCE: 6  
206 Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
207 1 5 10 15  
208 Ser Asp Lys Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
209 20 25 30  
210 Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
211 35 40 45  
212 Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
213 50 55 60  
214 Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
215 65 70 75 80  
216 Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
217 85 90 95  
218 Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
219 100 105 110  
220 Ala Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu  
221 115 120 125

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/581,976B

DATE: 11/26/2002  
TIME: 09:26:10

Input Set : A:\seqlist.txt  
Output Set: N:\CRF4\11262002\I581976B.raw

222 Cys Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val  
223 130 135 140  
224 Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe  
225 145 150 155 160  
226 Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Cys  
227 165 170 175  
228 Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr  
229 180 185 190  
230 Cys Tyr Ser Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro  
231 195 200 205  
232 Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys Pro Leu Cys  
233 210 215 220  
234 Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn  
235 225 230 235 240  
236 Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser  
237 245 250 255  
238 Arg Thr Arg Arg Glu Thr Gln Leu Met His Gly Asp Thr Pro Thr Leu  
239 260 265 270  
240 His Glu Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys  
241 275 280 285  
242 Tyr Glu Gln Leu Asn Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly  
243 290 295 300  
244 Pro Ala Gly Gln Ala Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr  
245 305 310 315 320  
246 Phe Cys Cys Lys Cys Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr  
247 325 330 335  
248 His Val Asp Ile Arg Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly  
249 340 345 350  
250 Ile Val Cys Pro Ile Cys Ser Gln Lys Pro Thr Ser Gly His His His  
251 355 360 365  
252 His His His  
253 370  
255 <210> SEQ ID NO: 7  
256 <211> LENGTH: 663  
257 <212> TYPE: DNA  
258 <213> ORGANISM: Artificial Sequence  
260 <220> FEATURE:  
261 <223> OTHER INFORMATION: Chimaeric protein (protein D from Haemophilus  
262 influenza B and mutated E7 from Human papilloma  
263 virus type 16)  
265 <400> SEQUENCE: 7  
266 atggatccaa gcagccatc atcaaatatg gcgaataccc aaatgaaatc agacaaaatc 60  
267 attattgctc accgtggtgc tagcggttat ttaccagagc atacgttaga atctaaagca 120  
268 cttgcgtttg cacaacaggc tgatttattt gagcaagatt tagcaatgac taaggatgg 180  
269 cgttagtgg ttattcacga tcactttt gatggcttga ctgatgtgc gaaaaaattc 240  
270 ccacatcgctc atcgtaaaga tggccgttac tatgtcatcg actttacctt aaaagaaaatt 300  
271 caaagtttag aatgacaga aaactttgaa accatggcca tgcatggaga tacacctaca 360  
272 ttgcataat atatgtttaga tttgcaacca gagacaactg atctctacgg ttatcagcaa 420  
273 tttaaatgaca gctcagagga ggaggatgaa atagatggtc cagctggaca agcagaaccg 480

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/581,976B

DATE: 11/26/2002

TIME: 09:26:11

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I581976B.raw